

Shifting machine learning for healthcare from development to deployment and from models to data

Year: 2022 | Citations: 300 | Authors: Angela Zhang, L. Xing, James Zou, Joseph C. Wu

Abstract

In the past decade, the application of machine learning (ML) to healthcare has helped drive the automation of physician tasks as well as enhancements in clinical capabilities and access to care. This progress has emphasized that, from model development to model deployment, data play central roles. In this Review, we provide a data-centric view of the innovations and challenges that are defining ML for healthcare. We discuss deep generative models and federated learning as strategies to augment datasets for improved model performance, as well as the use of the more recent transformer models for handling larger datasets and enhancing the modelling of clinical text. We also discuss data-focused problems in the deployment of ML, emphasizing the need to efficiently deliver data to ML models for timely clinical predictions and to account for natural data shifts that can deteriorate model performance. This Review discusses the use of deep generative models, federated learning and transformer models to address challenges in the deployment of machine learning for healthcare.